

TRANSMITTER LINEARIZATION

This application is a Continuation of International Application PCT/FI99/01051 filed December 16, 1999 which designated the U.S. and
 5 was published under PCT Article 21(2) in English.

BACKGROUND OF THE INVENTION

[0001] The invention relates to linearization of a radio transmitter.

[0002] The scarcity of radio frequencies leads to a need to use spectrum-efficient modulation methods in new radio systems. In Europe, a
 10 new radio system standard has been developed for PMR (Professional Mobile Radio) users, called TETRA (Terrestrial Trunked Radio). As the modulation method of the TETRA system, $\pi/4$ -DQPSK ($\pi/4$ -shifted Differential Quadrature Phase Shift Keying) has been chosen. From the transmitter point of view, a drawback of the modulation method is the
 15 variation in the amplitude of the envelope of the RF signal. In a nonlinear amplifier, such a variation causes InterModulation (IM). IM products spread the spectrum of the transmitted signal, and thus tend to reduce the benefits of using a linear modulation method. The IM products cannot usually be
 20 filtered as they form very close to the desired signal. With constant amplitude modulation methods, spectrum spreading does not occur, and consequently the signal may be amplified by a nonlinear amplifier.

[0003] A trunked PMR system, in which different user groups share same radio channels, has strict requirements regarding adjacent channel interference caused by a transmitter. These requirements call for a
 25 good linearity in the transmitter of the used radio system.

[0004] In a power amplifier, good linearity is only achieved with poor efficiency. However, the efficiency of portable devices should be as high as possible for the operation time to be adequate and in order not to waste battery capacity. Further, at least relatively good efficiency is required
 30 of the power amplifiers at base stations to avoid cooling problems. The achieving of adequate efficiency and linearity calls for linearizing the transmitter.

[0005] If the nonlinearities of an amplifier were known in advance, it would be possible to form inverse functions of the nonlinearities, and use
 35 them to convert the input signal, whereby the nonlinearities would be

cancelled. The properties of the amplifier do not, however, remain the same but vary due to ageing, warming up, and according to the radio channel and transmit power used. In addition, the amplifiers have individual differences. Need exists for linearization methods that must in an adaptive way be
5 capable of adapting to changing conditions.

[0006] Development work has been targeted into a number of different linearizing methods, and three of them have been found to possess qualities suitable for practical radio systems. These methods are feedforward, cartesian feedback and predistortion. A linearization method
10 can also be adaptive.

[0007] Thus, if the nonlinear transfer function of the amplifier is known and if it does not vary as a function of time, the signal to be transmitted can be linearized by applying a suitable transfer function causing a predistortion to the signal. This way, the signal coming out of the amplifier
15 can be made linear. This method is called predistortion. Predistortion is usually performed on baseband using a lookup table into which the conversion parameters causing the predistortion, i.e. the predistortion parameters, are stored.

[0008] For instance, with the temperature and age change of the amplifier its transfer function also changes and it becomes necessary to update the predistortion parameters in the lookup table. Prior art solutions use a gradient-based search for this purpose. A gradient-based search adapts towards minimizing an error. A problem in a gradient-based search is its slowness.
20

25 BRIEF DESCRIPTION OF THE INVENTION

[0009] It is thus an object of the invention to develop a method and an apparatus implementing the method so as to solve the above problems. The object of the invention is achieved by a method and a system characterized by what is stated in the independent claims 1 and 12.
30 Preferred embodiments of the invention are disclosed in the dependent claims.

[0010] The invention is based on the idea that an amplifier model is not formed and predistortion parameters are not defined using a gradient search, but the model is formed quickly and without a slow convergence on
35 the basis of a measured signal, i.e. a transmitter output signal, and a

corresponding input signal buffered into memory or any other corresponding ideal reference signal. The speed of the method of the invention is based on the fact that the nonlinearity of the transmitter is defined directly on the basis of a certain sample material in certain classes, such as classes according to the transmit power. This can be done faster than a definition based on minimizing the error, in which the nonlinearity of the transmitter must be determined by testing in different classes, whereby the number of tests may be large.

[0011] An advantage of the invention is that it provides a considerable improvement in prior art solutions by offering a faster way of updating the predistorter correction parameters. According to tests performed by the applicant, the method of the invention is up to 10 times faster than the gradient search used in prior art.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] The invention will now be described in greater detail in connection with preferred embodiments and with reference to the attached drawings in which

Figure 1 shows a block diagram of a transmitter of the invention according to one embodiment,

Figure 2 shows characteristic curves of the amplitudes of amplifiers and a predistorter, and

Figure 3 shows a characteristic curve of phase difference.

DETAILED DESCRIPTION OF THE INVENTION

[0013] Figure 1 shows a block diagram of a transmitter of the invention according to one embodiment. It should be noted that the figure shows only the elements that are essential for understanding the invention. In the figure, the transmitter receives an I and Q signal I_{IN} and Q_{IN} for transmission. Predistortion has been implemented using a digital signal processor (DSP) 1. Predistortion of the amplitude occurs in the multiplier elements 3A and 3B in accordance with an amplitude correction parameter obtained from an amplitude table 19. Correspondingly, a possible phase predistortion occurs in a phase shifter element 4 in accordance with a phase correction parameter obtained from a phase table 20. The predistorted signals are D/A-converted using D/A converters 5A and 5B into analogue signals which are also preferably filtered using low-pass reconstruction filters

6A and 6B. These signals are forwarded on to an I/Q modulator 7. The I/Q modulator 7 and an I/Q demodulator 15 operate according to the quadrature modulation principle. With this principle, it is possible to combine two independent signals in the transmitter, to transmit them on the same transmission band and then to separate them from each other in a receiver. The quadrature modulation principle is that two separate signals, I and Q (Inphase and Quadrature phase), are modulated using the same carrier frequency, but the phases of the carriers differ in that the carrier of the signal Q lags 90° behind the carrier of the signal I. After modulation, the signals are summed. Due to phase difference, the signals I and Q can be separated from each other when the sum signal is demodulated. In the I/Q modulator 7, the signals are modulated and combined. The I/Q modulator is synchronized by a local oscillator 10. The I/Q-modulated signal is forwarded to a power amplifier PA 8 and onward to an OUT antenna 2 for transmission. At the same time, feedback is formed by means of a sampling arrangement 9, for instance a directional coupler. A radio frequency (e.g. 400 MHz) feedback signal is preferably down-converted to an intermediate frequency of 450 kHz, for instance, using a down-converter 12. The down-conversion is synchronized by a local oscillator 11. The intermediate frequency signal can, if necessary, be filtered using a broadband filter 13 and subsequently attenuated. The intermediate frequency signal is sampled using an A/D converter 14 for baseband (or intermediate frequency) processing.

[0014] The I/Q demodulator 15 has been implemented using a digital signal processor 1. The A/D-converted intermediate frequency feedback signal is I/Q-demodulated with the I/Q demodulator 15 by digital multiplication into baseband I and Q signals. The I/Q demodulator is implemented by program in the digital signal processor 218, and a separate analogue I/Q demodulator is not needed. I/Q demodulation could also be performed as an analogue process without affecting the invention. The baseband I and Q feedback signals are preferably converted from an orthogonal presentation to a polar presentation in a converter 16B, whereby the amplitude and phase are directly obtained from the converted signals and forwarded to a calculation unit 17 in which the predistortion tables 19 and 20 are generated, which generation is described later in this description. The signals I_IN and Q_IN received by the transmitter for transmission,

which are preferably also converted to a polar presentation in the converter 16, are also forwarded to the calculation unit 17. The actual predistortion is done by means of the created predistortion tables 19 and 20. An absolute value of the complex signal formed by the signals I_IN and Q_IN, i.e. the amplitude of the signal, is defined in an element 18. This amplitude data is entered into the amplitude predistortion table 19 which provides a corresponding amplitude correction parameter to the multiplier elements 3A and 3B on the basis of the amplitude data. A corrected, i.e. predistorted amplitude data, formed using a multiplier element 21 is entered into the phase predistortion table 20. On the basis of this data, the phase table 20 provides the correct phase correction parameter to the phase shifter element 4. The amplitude data is preferably used as control data in the phase table 20, because a phase error in the transmitter depends on the amplitude of the signal. The amplitude data received by the phase table 20 is preferably also predistorted and thus corresponds to the amplitude of the signal coming out of the predistorter, whose magnitude determines the nonlinearity in the transmitter, and the phase predistortion required can be defined as accurately as possible. It is also possible to use amplitude predistortion only, which means that the phase predistortion table 20 and phase shifter element 4 are not needed.

[0015] The creation of the amplitude and phase predistortion tables 19 and 20 of the invention is performed as follows: samples 22 and 23 of the transmitter input signals I_IN and Q_IN and samples 24 and 25 of the (complex) transmitter output signal OUT are used as the inputs of the calculation element 17 in Figure 1. The signals 22 and 23 together form a complex reference signal REF in polar presentation. The reference signal REF is ideal in that it does not have any nonlinearity caused by the transmitter. Correspondingly, the signals 24 and 25 together form a complex feedback signal FB in polar presentation. The signals REF and FB are normalized so that the highest amplitude in both signals is 1. The reference signal REF and the corresponding feedback signal FB arrive at different times at the calculation element 17, because the transmitter causes a certain delay in the transmission of the signal, i.e. the value of the reference signal REF corresponding to a certain sample point of the incoming complex signal I_IN and Q_IN arrives at the calculation element 17 earlier than the value of the corresponding feedback signal FB. Because of this, the

reference signal REF is buffered in order to be able to compare in the calculation element 17 the value of a certain input complex signal I_IN and Q_IN which was input into the transmitter and the corresponding value which comes out of the transmitter. To form the predistortion tables 19 and 20, the predistortion parameters in tables 19 and 20 are set to values in which no predistortion occurs. After this, a suitable signal is fed into the transmitter and a predefined number of samples (e.g. 160 to 2,250) are taken from the signal I_IN and Q_IN coming in to the transmitter and from the signal OUT coming out of the transmitter. The amplitude values of the mutually corresponding sample points of the sample signals REF and FB are categorized into a required number of classes (e.g. 128 to 16,384) on the basis of the amplitude of the reference signal REF. The samples are preferably categorized on the basis of the amplitude, because the nonlinearity of the transmitter depends on the transmit power which, for its part, depends on the amplitude of the signal. Next, the values of the sample points of the feedback signal FB are compared with the corresponding buffered values of the reference signal REF, and the correction parameters are defined based on this comparison.

[0016] The comparison and the definition of the correction parameters are, for the part of the amplitude, illustrated in Figure 2. Straight line 31 shows the dependence of the normalized output amplitude of an ideal transmitter on the normalized input amplitude. Straight line 31 also shows the dependence of the reference signal REF on the input signal I_IN and Q_IN of the input signal, i.e. they are equal. Curve 32 shows the nonlinear characteristic curve of the transmitter (when no predistortion is used). The nonlinearity of the transmitter is mainly caused by the power amplifier 8. The transmitter can also have several power amplifiers 8 in series. Curve 32 is defined on the basis of the sample signals REF and FB for instance as follows: an average of the amplitudes of the sample points of the feedback signal FB in every class is calculated. Similarly, an average of the sample points of the reference signal REF corresponding to the sample points of the feedback signal FB in every class is calculated. The value of curve 32 at the centre point of the class in question in relation to the ideal curve 31 is obtained with the ratio of the calculated averages. In Figure 2, the centre points of classes are marked with circles. The ratio can also be calculated by first defining the ratio of the amplitudes of each sample point

of the feedback signal FB and the corresponding sample points of the reference signal REF and then defining the average of the ratios of the sample point pairs in each class. Curve 33 is the characteristic curve of the predistorter obtained by mirroring the characteristic curve 32 of the transmitter defined without predistortion in relation to the ideal curve 31. This is done, for instance, by defining a counterpart for each point (marked with circles) of the characteristic curve 32 of the transmitter. For instance, the counterpart of point 41 is determined by first finding a point of the ideal curve 31 that corresponds to the point 41 on the basis of the output amplitude value. In the example, where the output amplitude is 0.4, the result is the point 42 of the ideal curve 31. Next, a point is searched whose input amplitude value is the same as that of the defined ideal curve point 42 (0.4) and whose output amplitude value is the same as the input amplitude value (0.2) of the point 41 of the characteristic curve 32. In this case, the counterpart of the point 41 is the point 43. A counterpart is defined for each point of the characteristic curve 32 in the same way. The counterparts (marked with x in Figure 2) form the characteristic curve 33 of the predistorter. The characteristic curve 32 of the transmitter is discrete in relation to the input amplitude (the x axis), and the characteristic curve 33 of the predistorter defined on the basis of it is discrete in relation to the output amplitude (the y axis), i.e. they have been defined in the marked points only (the circles and the x's) due to the categorization of the samples into classes. The more classes are used, the more points are obtained, and, further, the closer one gets to the continuous curves 32 and 33 marked with a continuous line in the figure.

[0017] When the points of the predistorter characteristic curve 33 have been defined, the amplitude predistortion table 19 can be created. Because the predistorter characteristic curve 33 is discrete in relation to the output amplitude, i.e. the amplitude correction parameter has been defined for certain output amplitude values only, as described above, and because, on the other hand, the predistorter input amplitude can obtain any values, the correction parameter to be used for a certain input amplitude is defined by finding the point closest to the input amplitude on the discrete curve 33 and using the amplitude correction parameter corresponding to this point. The amplitude correction parameter refers here to a value by which the input amplitude should be multiplied to obtain the required output amplitude

according to the characteristic curve 33 of the predistorter. The correction parameter is thus obtained at a certain point of curve 33 by dividing the value of the output amplitude by the value of the input amplitude at the point in question. The amplitude predistortion table 19 can be formed by defining
 5 for a certain defined correction parameter a certain input amplitude value range within which the correction parameter is used. This can be done by means of the characteristic curve 33 by dividing curve 33 into ranges in relation to the input amplitude, the centre points of the ranges (marked with x's) being the points at which the output amplitude and correction parameter
 10 have been defined. The table can then be used to check, into which range the value of the input amplitude belongs, and the correction parameter corresponding to the range can then be used.

[0018] Figure 3 illustrates the definition of the correction parameters of the phase difference (between the sample point of the feedback signal FB and the sample point of the corresponding reference signal REF). The figure shows an example of phase difference as a function
 15 of a normalized amplitude (input amplitude) of the reference signal REF, i.e. curve 51. The characteristic curve 51 of the phase is formed by defining the average phase differences of the sample point pairs of the feedback signal FB and the reference signal REF in each class into which they have been categorized on the basis of the amplitude when defining the amplitude
 20 correction parameters as described above. In other words, the phase differences are categorized into classes on the basis of the normalized amplitude of the reference signal REF and the average of the phase differences is calculated for each class. This way, the characteristic curve
 25 values of the amplifier phase at the centre point of each class, i.e. the points marked with circles in Figure 3 forming the discrete characteristic curve of the phase, are obtained. The more classes are used, the smoother curve 51 becomes.

[0019] When the points of the characteristic curve 51 of the phase have been defined, the phase predistortion table 20 can be created. Because the characteristic curve 51 of the phase is discrete, i.e. the phase difference has been defined for certain input amplitude values (class
 30 averages) only, as described above, and, on the other hand, because the input amplitude can obtain any values, the correction parameter for a certain
 35 input amplitude is defined by finding the point closest to the input amplitude

on the discrete characteristic curve 51 of the phase and using the phase correction parameter corresponding to this point. The phase correction parameter refers here to the value which defines how much the phase of a signal arriving at the predistorter must be shifted and into which direction to achieve the required phase predistortion. The correction parameter is thus obtained at a certain point of the characteristic curve 51 by multiplying the phase difference value by -1 at the point in question, i.e. the phase is predistorted into the opposite direction by the amount of the phase difference. The phase predistortion table 20 can be formed, for instance, by defining for a certain defined correction parameter a certain input amplitude value range within which the correction parameter is used. This can be done by means of the characteristic curve 51 by dividing curve 51 into ranges in relation to the input amplitude, the centre points of the ranges (marked with circles) being the points at which the output amplitude and, thus also, the correction parameter have been defined. The table can then be used for checking, into which range the value of the input amplitude belongs, and the correction parameter corresponding to the range can then be used.

[0020] If a signal sample class does not contain any samples and the correction parameter can thus not be defined on the basis of the samples, a correction parameter of another class, preferably of a class closest to it, can be defined as the correction parameter of such a class. The correction parameter can also be defined by interpolation, or by some other method, on the basis of the correction parameters of several adjacent classes containing samples.

[0021] The format of the data in the predistortion tables 19 and 20 can differ from the above without making any difference to the basic idea of the invention. Similarly, the creation of the predistortion tables can differ from what is described above. Instead of a separate amplitude table 19 and phase table 20, also a combined predistortion table can be used in which a certain input amplitude value has a complex correction parameter which includes both the amplitude and phase correction parameters. It is also possible to use only amplitude predistortion, which reduces memory use, since a phase predistortion table is not needed.

[0022] To reiterate, the definition of the predistortion tables 19 and 20 is done as follows:

[0023] 1) The correction parameters in the amplitude predistortion table 19 are set to 1 (the signal is multiplied by 1, no predistortion occurs) and the correction parameters in the phase predistortion table 20 are set to 0 (the signal phase is shifted 0 degrees, no predistortion occurs).

5 [0024] 2) A normal modulated signal or another signal suited for training is transmitted.

[0025] 3) Samples are taken from the signal OUT being transmitted through the feedback branch.

[0026] 4) The delay caused by the circuit is compensated for.

10 [0027] 5) The corresponding samples 22 and 23 taken from the signals I_IN and Q_IN fed into the transmitter and the samples 24 and 25 taken from the signal OUT coming out of the transmitter are compared with each other and the characteristic curves of the transmitter are formed for the amplitude 33 and the phase 51.

15 [0028] 6) The predistortion tables 19 and 20 are created by means of the characteristic curves.

[0029] Although the use of the invention is here described mainly in connection with the TETRA system, this does not in any way limit the use of the invention in other types of systems. The structure of the transmitter used may differ from what is described herein without deviating from the basic idea of the invention.

20 [0030] It is obvious to a person skilled in the art that while technology advances, the basic idea of the invention may be implemented in many different ways. The invention and its embodiments are thus not limited to the examples described above, but can vary within the scope of the claims.